RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: _10/535,667_
Source: _1Fwg_
Date Processed by STIC: _3/6/06_

ENTERED

IFWO

RAW SEQUENCE LISTING   DATE: 03/06/2006
PATENT APPLICATION: US/10/535,667   TIME: 12:01:48

Input Set : **A:\PTO.TS.txt**
Output Set: **N:\CRF4\03062006\J535667.raw**

```
 3 <110> APPLICANT: IMAMURA, Hiromi
 4        YOSHIDA, Masasuke
 5        YOKOYAMA, Ken
 7 <120> TITLE OF INVENTION: A Rotary Motor molecule V1-ATPase
 9 <130> FILE REFERENCE: 2005_0822A/WMC/00653
11 <140> CURRENT APPLICATION NUMBER: US 10/535,667
12 <141> CURRENT FILING DATE: 2005-05-19
14 <150> PRIOR APPLICATION NUMBER: JP2002-337212
15 <151> PRIOR FILING DATE: 2002-11-20
17 <150> PRIOR APPLICATION NUMBER: PCT/JP2003/012983
18 <151> PRIOR FILING DATE: 2003-10-09
20 <160> NUMBER OF SEQ ID NOS: 5
22 <170> SOFTWARE: PatentIn Ver. 2.1
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 4199
26 <212> TYPE: DNA
27 <213> ORGANISM: Thermus thermophilus
29 <220> FEATURE:
30 <221> NAME/KEY: CDS
31 <222> LOCATION: (1)..(318)
33 <220> FEATURE:
34 <221> NAME/KEY: CDS
35 <222> LOCATION: (334)..(2067)
37 <220> FEATURE:
38 <221> NAME/KEY: CDS
39 <222> LOCATION: (2081)..(3514)
41 <220> FEATURE:
42 <221> NAME/KEY: CDS
43 <222> LOCATION: (3528)..(4196)
45 <400> SEQUENCE: 1
46 gtg agg atg gcg gtg atc gcc gat ccc gag acc gcc cag ggg ttc cgg    48
47 Val Arg Met Ala Val Ile Ala Asp Pro Glu Thr Ala Gln Gly Phe Arg
48   1               5                   10                  15
50 ctc gcg ggc ctc gag ggc tac ggg gcc tct tcg gcg gag gag gcc caa    96
51 Leu Ala Gly Leu Glu Gly Tyr Gly Ala Ser Ser Ala Glu Glu Ala Gln
52               20                  25                  30
54 agc ctc ctg gaa acc ctc gtg gag cgg ggc ggc tac gcc ctg gtg gcc   144
55 Ser Leu Leu Glu Thr Leu Val Glu Arg Gly Gly Tyr Ala Leu Val Ala
56           35                  40                  45
58 gtg gac gag gcg ctc ctc ccc gac ccc gag cgg gcg gtg gag cgc ctc   192
59 Val Asp Glu Ala Leu Leu Pro Asp Pro Glu Arg Ala Val Glu Arg Leu
60       50                  55                  60
62 atg cgg ggc agg gac ctc ccc gtg ctc ctg ccc atc gcg ggg ctg aag   240
```

```
63 Met Arg Gly Arg Asp Leu Pro Val Leu Leu Pro Ile Ala Gly Leu Lys
64  65                    70                75                80
66 gag gcc ttc cag ggg cac gac gtg gaa ggc tac atg cgg gag ctg gtg   288
67 Glu Ala Phe Gln Gly His Asp Val Glu Gly Tyr Met Arg Glu Leu Val
68              85                90                95
70 agg aag acc atc ggc ttt gac atc aag ctg tagaatggag ggacg atg atc   339
71 Arg Lys Thr Ile Gly Phe Asp Ile Lys Leu                    Met Ile
72          100             105
74 caa ggg gtg atc cag aag atc gcg ggc ccg gcg gtg atc gcc aag ggc   387
75 Gln Gly Val Ile Gln Lys Ile Ala Gly Pro Ala Val Ile Ala Lys Gly
76     110             115             120
78 atg ctc ggg gcc cgc atg tac gac atc tgc aag gtg ggc gaa gag ggc   435
79 Met Leu Gly Ala Arg Met Tyr Asp Ile Cys Lys Val Gly Glu Glu Gly
80 125             130             135             140
82 ctc gtg ggc gag atc atc cgc ctg gac ggg gac acg gcc ttc gtc cag   483
83 Leu Val Gly Glu Ile Ile Arg Leu Asp Gly Asp Thr Ala Phe Val Gln
84              145             150             155
86 gtc tac gag gac acc tcg ggc cta aag gtg ggg gag ccc gtg gtc tcc   531
87 Val Tyr Glu Asp Thr Ser Gly Leu Lys Val Gly Glu Pro Val Val Ser
88              160             165             170
90 acg ggg ctt ccc ttg gcg gtg gag ctc ggc ccc ggg atg ctg aac ggc   579
91 Thr Gly Leu Pro Leu Ala Val Glu Leu Gly Pro Gly Met Leu Asn Gly
92          175             180             185
94 atc tac gac ggc atc cag cgc ccc ctg gag cgc atc cgg gag aag acg   627
95 Ile Tyr Asp Gly Ile Gln Arg Pro Leu Glu Arg Ile Arg Glu Lys Thr
96     190             195             200
98 ggg atc tac atc acc cgg ggc gtg gtg gtc cac gcc ctg gac cgg gag   675
99 Gly Ile Tyr Ile Thr Arg Gly Val Val Val His Ala Leu Asp Arg Glu
100 205            210             215             220
102 aag aag tgg gcc tgg acg ccc atg gtc aag ccc ggg gac gag gtg cgg   723
103 Lys Lys Trp Ala Trp Thr Pro Met Val Lys Pro Gly Asp Glu Val Arg
104             225             230             235
106 ggg ggt atg gtc ctg ggc acg gtg ccc gag ttc ggc ttc acc cac aag   771
107 Gly Gly Met Val Leu Gly Thr Val Pro Glu Phe Gly Phe Thr His Lys
108             240             245             250
110 atc ctg gta ccc ccg gac gtg cgg ggc cgg gtc aag gag gtg aag ccc   819
111 Ile Leu Val Pro Pro Asp Val Arg Gly Arg Val Lys Glu Val Lys Pro
112         255             260             265
114 gcc ggg gag tac acc gtg gag gag ccg gtg gtg gtc ctc gag gac ggc   867
115 Ala Gly Glu Tyr Thr Val Glu Glu Pro Val Val Val Leu Glu Asp Gly
116     270             275             280
118 acc gag ctc aag atg tac cac acc tgg ccc gtt cgc cgg gcg agg ccc   915
119 Thr Glu Leu Lys Met Tyr His Thr Trp Pro Val Arg Arg Ala Arg Pro
120 285             290             295             300
122 gtg caa agg aag ctt gac ccc aac acc ccc ttc ctc acg ggg atg cgc   963
123 Val Gln Arg Lys Leu Asp Pro Asn Thr Pro Phe Leu Thr Gly Met Arg
124             305             310             315
126 atc ctg gac gtc ctc ttc ccc gtg gcc atg ggg ggc acc gcc gcc atc   1011
127 Ile Leu Asp Val Leu Phe Pro Val Ala Met Gly Gly Thr Ala Ala Ile
```

```
128                 320                 325                 330
130 cct ggg ccc ttc ggc agc ggc aag acc gtg acc cag cag tcc ctg gcc    1059
131 Pro Gly Pro Phe Gly Ser Gly Lys Thr Val Thr Gln Gln Ser Leu Ala
132         335                 340                 345
134 aag tgg tcc aac gcc gac gtg gtg gtc tac gtg ggc tgc ggg gag cgg    1107
135 Lys Trp Ser Asn Ala Asp Val Val Val Tyr Val Gly Cys Gly Glu Arg
136         350                 355                 360
138 ggg aac gag atg acc gac gtg ctc gtg gag ttc ccc gag ctc acc gac    1155
139 Gly Asn Glu Met Thr Asp Val Leu Val Glu Phe Pro Glu Leu Thr Asp
140 365                 370                 375                 380
142 ccc aag acg ggg ggg ccc ttg atg cac cgc acc gtc ctc atc gcc aac    1203
143 Pro Lys Thr Gly Gly Pro Leu Met His Arg Thr Val Leu Ile Ala Asn
144                 385                 390                 395
146 acc tcc aac atg ccc gtg gcc gcc cgc gag gcc agc atc tac gtg ggc    1251
147 Thr Ser Asn Met Pro Val Ala Ala Arg Glu Ala Ser Ile Tyr Val Gly
148                 400                 405                 410
150 gtg acc atc gcc gag tac ttc cgc gac cag ggc ttc tcc gtg gcc ctc    1299
151 Val Thr Ile Ala Glu Tyr Phe Arg Asp Gln Gly Phe Ser Val Ala Leu
152         415                 420                 425
154 atg gcc gac tcc acg agc cgc tgg gcc gag gct ttg cgc gag atc tct    1347
155 Met Ala Asp Ser Thr Ser Arg Trp Ala Glu Ala Leu Arg Glu Ile Ser
156         430                 435                 440
158 agc cgc ctc gag gag atg ccc gcc gag gag ggc tac ccg ccc tac ctc    1395
159 Ser Arg Leu Glu Glu Met Pro Ala Glu Glu Gly Tyr Pro Pro Tyr Leu
160 445                 450                 455                 460
162 gcc gcc agg ctc gcc gcc ttc tac gag cgg gcg ggc aag gtc atc acc    1443
163 Ala Ala Arg Leu Ala Ala Phe Tyr Glu Arg Ala Gly Lys Val Ile Thr
164                 465                 470                 475
166 ctg ggc ggc gag gag ggg gcg gtg acc atc gtg ggg gcc gtc tcc ccg    1491
167 Leu Gly Gly Glu Glu Gly Ala Val Thr Ile Val Gly Ala Val Ser Pro
168                 480                 485                 490
170 ccg ggc ggc gac atg tcc gag ccc gtg acc cag tcc acc ttg agg atc    1539
171 Pro Gly Gly Asp Met Ser Glu Pro Val Thr Gln Ser Thr Leu Arg Ile
172         495                 500                 505
174 gtg ggg gcc ttc tgg cgg ctt gac gcc tcc ctg gcc ttc cgc cgc cac    1587
175 Val Gly Ala Phe Trp Arg Leu Asp Ala Ser Leu Ala Phe Arg Arg His
176 510                 515                 520
178 ttc ccc gcc atc aac tgg aac ggc tcc tac agc ctc ttc acc tcc gcc    1635
179 Phe Pro Ala Ile Asn Trp Asn Gly Ser Tyr Ser Leu Phe Thr Ser Ala
180 525                 530                 535                 540
182 ctt gac ccc tgg tac cgg gag aac gtg gcc gag gac tac ccc gag ctc    1683
183 Leu Asp Pro Trp Tyr Arg Glu Asn Val Ala Glu Asp Tyr Pro Glu Leu
184                 545                 550                 555
186 cgc gac gcc atc tcc gag ctt ttg cag cgg gag gcg ggc ctc cag gag    1731
187 Arg Asp Ala Ile Ser Glu Leu Leu Gln Arg Glu Ala Gly Leu Gln Glu
188                 560                 565                 570
190 atc gtc cag ctc gtg ggg ccg gac gcc ctc cag gac gcc gag cgc ctc    1779
191 Ile Val Gln Leu Val Gly Pro Asp Ala Leu Gln Asp Ala Glu Arg Leu
192         575                 580                 585
```

RAW SEQUENCE LISTING   DATE: 03/06/2006
PATENT APPLICATION: US/10/535,667   TIME: 12:01:48

Input Set : A:\PTO.TS.txt
Output Set: N:\CRF4\03062006\J535667.raw

```
194 gtc att gag gtg ggc cgg atc atc cgc gag gac ttc ctg cag cag aac   1827
195 Val Ile Glu Val Gly Arg Ile Ile Arg Glu Asp Phe Leu Gln Gln Asn
196     590                 595                 600
198 gcc tac cac gag gtg gac gcc tac tgc tcc atg aag aag gcc tac ggg   1875
199 Ala Tyr His Glu Val Asp Ala Tyr Cys Ser Met Lys Lys Ala Tyr Gly
200 605                 610                 615                 620
202 atc atg aag atg atc ctc gcc ttc tac aag gag gcg gag gcg gcc atc   1923
203 Ile Met Lys Met Ile Leu Ala Phe Tyr Lys Glu Ala Glu Ala Ala Ile
204                 625                 630                 635
206 aag cgg ggg gtt tcc ata gac gag atc ctg cag ctc ccc gtt ctg gag   1971
207 Lys Arg Gly Val Ser Ile Asp Glu Ile Leu Gln Leu Pro Val Leu Glu
208                 640                 645                 650
210 cgc atc ggc cgc gcc cgc tac gtg agc gag gag gag ttc ccc gcc tac   2019
211 Arg Ile Gly Arg Ala Arg Tyr Val Ser Glu Glu Glu Phe Pro Ala Tyr
212         655                 660                 665
214 ttt gag gag gcc atg aag gag atc cag ggg gcc ttc aag gcc ctg gcc   2067
215 Phe Glu Glu Ala Met Lys Glu Ile Gln Gly Ala Phe Lys Ala Leu Ala
216     670                 675                 680
218 taaagggggga gag atg gac ctt ctg aag aag gac tac acg ggc atc acc   2116
219             Met Asp Leu Leu Lys Lys Glu Tyr Thr Gly Ile Thr
220                 685                 690                 695
222 tac atc tcg ggg cct ctt ctc ttc gtg gag aac gcc aag gac ctg gcc   2164
223 Tyr Ile Ser Gly Pro Leu Leu Phe Val Glu Asn Ala Lys Asp Leu Ala
224                 700                 705                 710
226 tac ggg gcc atc gtg gac atc aag gac ggc acg ggc cgg gtc cgc ggc   2212
227 Tyr Gly Ala Ile Val Asp Ile Lys Asp Gly Thr Gly Arg Val Arg Gly
228         715                 720                 725
230 ggc cag gtg att gag gtc tcc gag gag tac gcc gtc atc cag gtg ttt   2260
231 Gly Gln Val Ile Glu Val Ser Glu Glu Tyr Ala Val Ile Gln Val Phe
232     730                 735                 740
234 gag gaa acc act ggg ctg gac ctg gcc acg acc agc gtg agc ctg gtg   2308
235 Glu Glu Thr Thr Gly Leu Asp Leu Ala Thr Thr Ser Val Ser Leu Val
236 745                 750                 755                 760
238 gag gac gtg gcc cgg ctt ggg gtc tcc aag gag atg ctg ggc cgc cgc   2356
239 Glu Asp Val Ala Arg Leu Gly Val Ser Lys Glu Met Leu Gly Arg Arg
240                 765                 770                 775
242 ttc aac ggc atc ggc aag ccc ata gac ggc ctg ccg ccc atc acc ccg   2404
243 Phe Asn Gly Ile Gly Lys Pro Ile Asp Gly Leu Pro Pro Ile Thr Pro
244                 780                 785                 790
246 gag aag cgg ctc ccc atc acc ggc ctt ccc tta aac ccc gtg gcc cgg   2452
247 Glu Lys Arg Leu Pro Ile Thr Gly Leu Pro Leu Asn Pro Val Ala Arg
248                 795                 800                 805
250 agg aag ccg gag cag ttc atc cag acg ggc atc tcc acc att gac gtg   2500
251 Arg Lys Pro Glu Gln Phe Ile Gln Thr Gly Ile Ser Thr Ile Asp Val
252     810                 815                 820
254 atg aac acc ctg gtc cgg ggg cag aag ctt ccc atc ttc tcc ggc tcg   2548
255 Met Asn Thr Leu Val Arg Gly Gln Lys Leu Pro Ile Phe Ser Gly Ser
256 825                 830                 835                 840
258 ggg ctt ccc gcc aac gag atc gcc gcc cag atc gcc cgc cag gcc acg   2596
```

```
259 Gly Leu Pro Ala Asn Glu Ile Ala Ala Gln Ile Ala Arg Gln Ala Thr
260             845             850             855
262 gtg cgc ccc gac ctc tcc ggg gag ggg gag aag gag gag ccc ttc gcc   2644
263 Val Arg Pro Asp Leu Ser Gly Glu Gly Glu Lys Glu Glu Pro Phe Ala
264             860             865             870
266 gtg gtc ttc gcc gcc atg ggg atc acg cag cgg gag ctc tcc tac ttc   2692
267 Val Val Phe Ala Ala Met Gly Ile Thr Gln Arg Glu Leu Ser Tyr Phe
268         875             880             885
270 atc cag gag ttt gag cgc acc ggg gcc ctg agc cgc tcc gtc ctc ttc   2740
271 Ile Gln Glu Phe Glu Arg Thr Gly Ala Leu Ser Arg Ser Val Leu Phe
272     890             895             900
274 ctg aac aag gcg gac gac ccc acc att gag cgc atc ctc acc ccc cgc   2788
275 Leu Asn Lys Ala Asp Asp Pro Thr Ile Glu Arg Ile Leu Thr Pro Arg
276 905             910             915             920
278 atg gcc ctc acc gtg gcc gag tac ctg gcc ttt gag cac gac tac cac   2836
279 Met Ala Leu Thr Val Ala Glu Tyr Leu Ala Phe Glu His Asp Tyr His
280             925             930             935
282 gtc ctc gtc atc ctc acg gac atg acc aac tac tgc gag gcc ttg cgg   2884
283 Val Leu Val Ile Leu Thr Asp Met Thr Asn Tyr Cys Glu Ala Leu Arg
284             940             945             950
286 gag atc ggg gcc gcc cgc gag gag atc ccg ggc cgc cgc ggt tac ccc   2932
287 Glu Ile Gly Ala Ala Arg Glu Glu Ile Pro Gly Arg Arg Gly Tyr Pro
288         955             960             965
290 ggc tac atg tac acc gac ctg gcc acc atc tac gag cgc gcc ggg gtg   2980
291 Gly Tyr Met Tyr Thr Asp Leu Ala Thr Ile Tyr Glu Arg Ala Gly Val
292     970             975             980
294 gtg gag ggg aag aag ggg agc gtg acc cag atc ccc atc ctc tcc atg   3028
295 Val Glu Gly Lys Lys Gly Ser Val Thr Gln Ile Pro Ile Leu Ser Met
296 985             990             995             1000
298 ccc gac gac gac cgc acc cac ccc atc ccc gac ctc acg ggc tac atc   3076
299 Pro Asp Asp Asp Arg Thr His Pro Ile Pro Asp Leu Thr Gly Tyr Ile
300             1005            1010            1015
302 acc gag ggg cag atc cag ctc tcc cgg gag ctc cac cgc aag ggc atc   3124
303 Thr Glu Gly Gln Ile Gln Leu Ser Arg Glu Leu His Arg Lys Gly Ile
304         1020            1025            1030
306 tac ccg ccc att gac ccc ttg ccc tcc ctc tcc cgg ctc atg aac aac   3172
307 Tyr Pro Pro Ile Asp Pro Leu Pro Ser Leu Ser Arg Leu Met Asn Asn
308         1035            1040            1045
310 ggc gtg ggc aag ggc aag acc cgg gag gac cac aag cag gtc tcc gac   3220
311 Gly Val Gly Lys Gly Lys Thr Arg Glu Asp His Lys Gln Val Ser Asp
312     1050            1055            1060
314 cag ctc tac tcc gcc tac gcc aac ggg gtg gac atc cgg aag ctc gtg   3268
315 Gln Leu Tyr Ser Ala Tyr Ala Asn Gly Val Asp Ile Arg Lys Leu Val
316 1065            1070            1075            1080
318 gcc atc atc ggc gag gac gcc ctc acg gag aac gac cgc cgt tac ctc   3316
319 Ala Ile Ile Gly Glu Asp Ala Leu Thr Glu Asn Asp Arg Arg Tyr Leu
320             1085            1090            1095
322 cag ttc gcc gac gcc ttt gaa cgg ttc ttc atc aac cag ggg cag cag   3364
323 Gln Phe Ala Asp Ala Phe Glu Arg Phe Phe Ile Asn Gln Gly Gln Gln
```

VERIFICATION SUMMARY   DATE: 03/06/2006
PATENT APPLICATION: US/10/535,667   TIME: 12:01:49

Input Set : **A:\PTO.TS.txt**
Output Set: **N:\CRF4\03062006\J535667.raw**